Journal of Agriculture and Technology Management (JATM)
23(1): 9-13 (January-June 2020)

Original article

Establishing the phylogenetic relationship of the capsicum genus based on the rpl20 gene

Jleian Mard M. Loseñara[1]*, and Lance Ylanan[2]

[1]College of Arts and Sciences at Tuburan Campus, Cebu Technological University, Barangay VIII, Tuburan, 6043, Philippines
[2]Department of Biology at Talamban Campus, University of San Carlos, Talamban, Cebu, 6000, Philippines

ABSTRACT

Extremely diverse and widespread, the Solanaceae family is of great economic importance, especially one of its genera, *Capsicum.* Globally, the production of sweet and chili pepper crops is an important agricultural commodity. Owing to its complex taxonomy, morphological classification, usually based on flowers, is not enough. With the advent of molecular analysis, it is beneficial to create a phylogenetic tree based on genetic data to trace the genus genus' evolutionary relationship, which could serve as baseline work for future studies on the Capsicum genus along with morphological data. In this study, five divergent genes were identified, with rpl20 being the most divergent. A phylogenetic tree generated from Mega X based on the chloroplast plastome rpl20 revealed *C. lycianthoides* as an outgroup to the other Capsicum species. In addition, *C. galapagoense* and *C. tovarii* as more evolutionary diverged species were inferred based on the findings.

KEYWORDS: *Capsicum, phylogeny, rpl20, phylogenetic relationship*

1 INTRODUCTION

The Solanaceae family is enormously widespread and diverse, with an estimated 85-90 genera and an approximate 1400-1700 species, rendering it one of the largest families of flowering plants (Woodland, 1997). Its members are extremely diverse; that ranges from trees to small annual herbs; in terms of habitat, deserts to the rainiest tropical rain forests; and in physical characters, with notable variation in many morphologies of both flowers and fruits (Knapp 2002). A significant number of common crop plants are members of the family, including the tomato (*Lycoperiscon esculentum*), chilli peppers (Capsicum spp.) aubergine (*Solanum melongena, S. aethiopicum*), and husk tomato (*Physalis* spp., about 100 species), the ornamental garden plant termed as "Chinese lanterns." Several poisonous plants in the family include the deadly nightshade (*Atropa belladonna*), bittersweet (*Solanum dulcamara*), henbane (*Hyoscyamus niger*), and thornapple (*Datura stramonium*) while some, *Petunia spp*., *Brugmansia spp.* (jungle trumpets) and *Nicotiana spp.* are reared as garden ornamentals (Passarin *et al,* 2008). Both wild varieties and cultivates of the peppers and chilies, members of the *Capsicum* genus, are of significant economic value (D'Agostino *et al,* 2018).

An important agribusiness worldwide, the production of sweet and chili pepper crops is one where the market stimulates family farming and increased employment and income generation from agriculture (Reifschneider and Ribeiro, 2008). According to the Food and Agricultural Organization of the United Nations (FAOSTAT, 2010), Asia is the leading producer where much of the production is generated by China, followed by India. Owing to the selection process, (Clement *et al,* 2010), varieties with novel morphological forms stem in such areas whereby their genetic variation is not clearly identified. Generally, species and genus identification is performed based on morphological characters or features, mainly in the flowers.

Over the past decade, sequencing of genomes has been an efficient method to investigate and elucidate a wide range of biological systems. With the use of next-generation sequencing technologies, comparative and evolutionary genomic studies such as in higher plants have been boosted. Extensively utilized to reconstruct the phylogeny of related species are the chloroplast genes. A number of chloroplast regions, including intergenic spaces and genes (Passarin *et al,* 2008), have been analyzed to obtain more phylogenetic reconstruction information. According to Chiang and Schaal (2000), the plastome (plastid genome) is regarded as an integral tool in plant phylogeny studies, especially in the above species level owing to both low-frequency changes in the chloroplast DNA molecule (cpDNA) and low rate of sequence evolution. Along with a few others, the chloroplast ribulose biphosphate carboxylase large (rbcL) gene sequences have been extensively utilized to establish phylogenetic relationships in plants (Miz *et al,* 2008).

*corresponding author: Jleianmard.losenara@ctu.edu.ph
p-ISSN: 2599-4875 e-ISSN: 2599-4980

However, in some cases, the phylogenetic data contained in a single chloroplast genetic marker does not have enough resolving power to distinguish between close taxa, especially in low taxonomic levels. Increasing the number of characters consequently results to improving the resolution of phylogenetic analysis. Because the chloroplast genome does not commonly go through systematic recombination, each of the various markers in chloroplasts can be put together in an analysis belonging to the same haplotype.

Grounded on several taxonomic studies of the *Capsicum* genus, Passarin *et al,* (2018) grouped Capsicum species into three main complexes*: C. annuum* (CA), *C. baccatum* (CB), and *C. pubescens* (CP) based on morphological, biochemical, cytogenetic and ethnobotanical data. The CA complex includes both domesticated and wild species of *C. annuum, Capsicum chinense*, *Capsicum frutescens*, and *Capsicum galapagoense*; the CB complex includes *C. baccatum*, *Capsicum praetermissum*, and *Capsicum tovarii*, whereas the CP complex is comprised by *C. pubescens*, *Capsicum eximium*, and *Capsicum cardenasii*. Considered a bridge species, *Capsicum chacoense* could be included in either the CA or CB complex. Though this classification is commonly accepted, establishing the genetic relationships between and within the *Capsicum* genus is still debated. However, several studies have been conducted utilizing different nuclear and plastid molecular markers to obtain a better insight of the genetic relationships within *Capsicum*.

This paper aimed to establish the phylogenetic relationship of *Capsicum* species based on molecular data. Specifically, this study aims to identify an appropriate gene and its sequence that could be utilized as a basis to determine phylogenetic relationship of ten *Capsicum* species. In addition, using the identified gene sequence, this paper intends to generate a phylogenetic tree showing the evolutionary relationship of the different Capsicum species.

2 MATERIALS AND METHODS

To determine the phylogenetic relationship of ten identified *Capsicum* species, a conserved gene was identified based on available existing literature. The nucleotide sequences of the rpl20 gene of the *Capsicum* species were obtained from GenBank Database by gaining access to the National Center for Biotechnology Information (NCBI) website. Providing access to genetic and biomedical information, the NCBI website (https://www.ncbi.nlm.nih.gov/) can be accessed online. The gene sequences were inputted to the Mega X software, short for Molecular Evolutionary Genetic Analysis, a free software that is an integrated tool for conducting statistical analysis of molecular evolution and constructing phylogenetic trees based on gene sequence. MEGA consists of an enormous repertoire of programs, including assembly of sequence alignments, inferring evolutionary trees, estimating genetic distances and diversities, computing time trees, and inferring ancestral sequences (Kumar *et al,* 2016).

3 RESULTS AND DISCUSSIONS

A study conducted by D'Agostino *et al,* (2018) established the complete nucleotide sequences of plastomes of eleven genotypes belonging to eight *Capsicum* species representing the three main taxonomic complexes. In addition, the study conducted a genome-wide analysis of molecular diversity among Capsicum plastomes was also performed. From eight Capsicum species, a group of eleven genotypes representing three complexes of the genus Capsicum was sampled for chloroplast isolation, cpDNA extraction, and sequencing. For the CA complex, the researchers sampled one genotype each for the species *C. chinense* (chi), *C. frutescens* (fru), *C. galapagoense* (gal) and three *C. annuum* genotypes (ann1, ann2, and ann3) and. For the CB complex, we sampled C. *baccatum subsp. Baccatum* (bac.b), *C. baccatum subsp. pendulum* (bac.p), and *C. praetermissum* (pra). Finally, the researchers also included a genotype from the CP complex, namely *C. chacoense* genotype (cha) and *C. pubescens* (pub) that, depending on the classification schemes, is included in either CA or in CB.

Figure 1, taken from the study of D'Agostino *et al,* (2018), shows the results of the analysis of molecular evolution of plastid genes within the *Capsicum* genus. The first column of figure 1, marked A, illustrates the estimation of protein-coding gene divergence by the average branch length and standard deviation for each gene tree. In addition, the second column, B, depicts the number of putative sites under positive selection as inferred by the study, gene divergence analysis, based on Selection, and indicated that protein-coding genes are generally well conserved among *Capsicum* species.

The most divergent genes are rpl20 and rpl32, followed by rpl36, clpP, and accD. Basing on the figure, rpl20 appeared to be the most divergent gene. Additionally, Jo *et al,* (2011) stated that in particular, rpl20 is one of two that demonstrated the highest variability between species as already observed upon comparing the *C. annuum* plastome with those of other Solanaceae.

Table 1 shows the identified *Capsicum* species reflecting the chloroplast genomes which were released into the public domain with accession numbers. The data is downloaded from GenBank. D'Agostino *et al,* (2018) looked into the coding sequences of the 79 protein-coding genes present in all *Capsicum* plastomes and in *C. lycianthoides* (NC_026551). These were extracted and fed into the Selecton web server

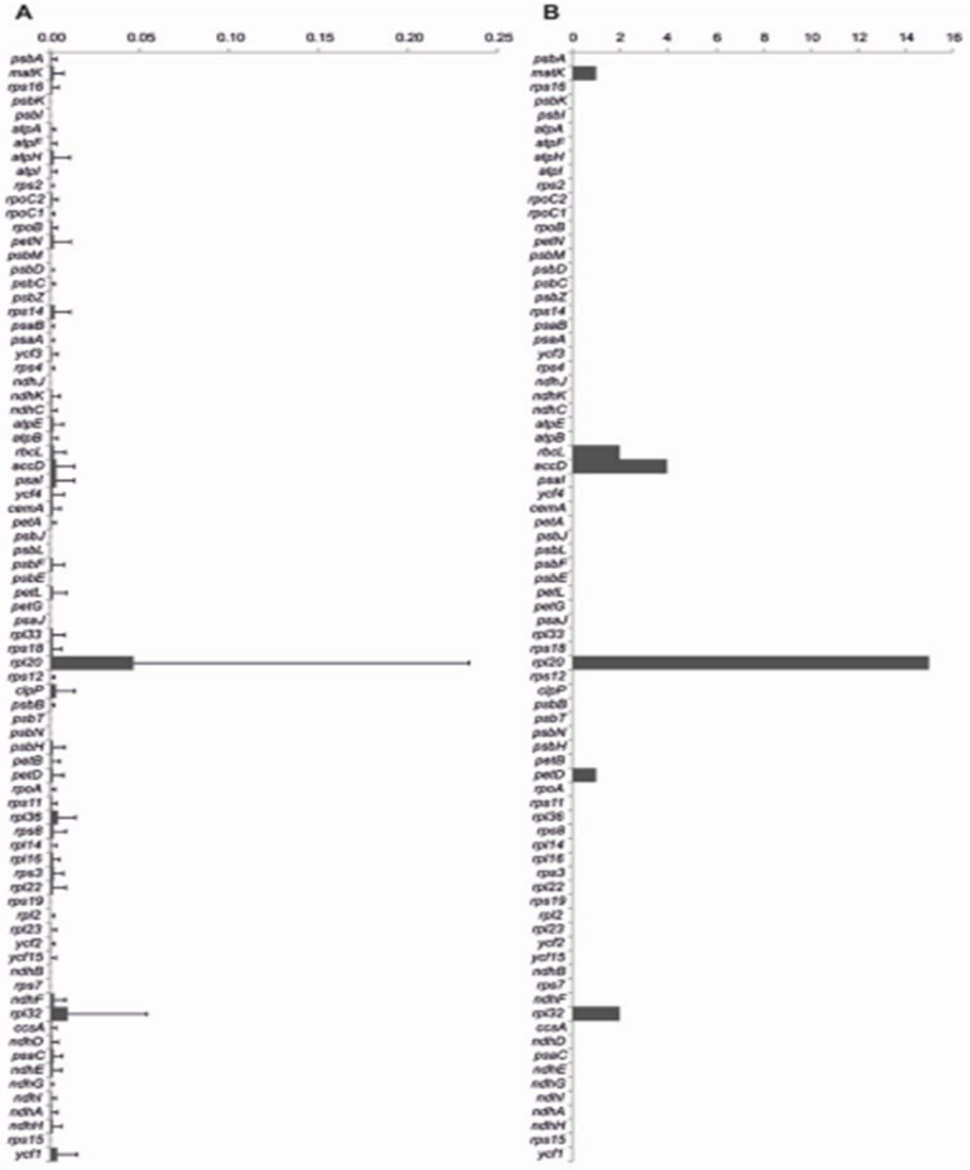

Figure 1. Results of Molecular Evolution Analysis of Plastid Genes within the Capsicum Genus adapted from D'Agostino *et al,* 2018.

Table 1. rpl20 Chloroplast Genomic Sequences and Accession Numbers of *Capsicum* Species

Species	***rpl20* Chloroplast Genomic Sequence**	**Accession Number**
C. frutescens	NC_028007.1	KR078312.1
C. chacoense	NC_033525.1	KX913218.1
C. chinense	NC_030543.1	KU041709.1
C. galapagoense	NC_033524.1	KX913216.1
C. annuum	NC_018552.1	JX270811.1
C. tovarii	NC_033526.1	KX913219.1
C. baccatum var. baccatum	NC_039693.1	KR078314.1
C. pubescens	NC_039694.1	MH559325.1
C. lycianthoides	NC_026551.1	KP274856.1
C. eximium	NC_033527.1	KX913220.1

(http://selecton.tau.ac.il/) in order to examine amino-acid sites under positive selection. As revealed by the study of D'Agostino *et al,* (2018), along with the annotated 79 protein-coding genes, 26 have perfectly conserved sequences and 48 have point mutations within the coding sequence, while five genes, namely accD, ndhB, rpl20, ycf1, and ycf2 are the most variable.

Figure 2 illustrates the evolutionary relationship of the ten *capsicum* species based on the rpl20 gene sequence through a phylogenetic tree. Currently, plastomes of only 10 *Capsicum* species are available in the database, thus, only these species can be included in constructing the tree. Using the Neighbor-Joining method (Saitou *et al,* 1987), the evolutionary history was inferred. With the sum of branch length = 20.00000000, the optimal tree is shown. To infer the phylogenetic tree, branch lengths were in the same units as those of the evolutionary distances used, the tree is drawn to scale. Computed using the number of base differences method (Nei, 2000), the evolutionary distances were in number of base differences per sequence as units. The analysis involved 10 nucleotide sequences. All gap-containing positions and missing data were eliminated. There were a total 404 positions in the final dataset.

The tree places *C. lycianthoides* as an outgroup to the other *Capsicum* species. *C. chacoense* and *C. pubescens* are sister group to *C. lycianthoides*. The phylogenetic tree generated by the MEGA X software places *C. annuum* as sister to group to more divergent species including *C. chinense, C. eximium, C. frutescens.*

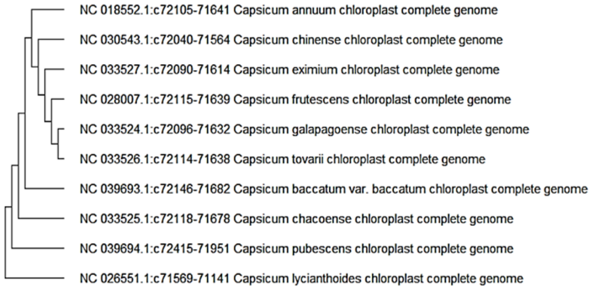

Figure 2. Evolutionary Relationship of Capsicum Species based on rpl20 Plastid Gene

The tree also indicates that *C. galapagoense* and *C. tovarii* are more evolutionary diverged species. *Capsicum galapagoense* is an endemic species in the Galápagos Islands off the coast of Ecuador in South America. This species is endemic and is a rather small plant with distinctive, densely hairy leaves which spread strong, pleasant odor when touched. A new species of Capsicum, *Capsicum tovarii* is found on the slopes of the Mantaro river basin, in Huancavelica and Ayacucho of Peru. To date, one of the most interesting aspects of *C. tovarii* is the peculiar nature of its breeding system, which is under continuing investigation. Eshbaugh *et al,* (1983) have noted that certain individuals function only as male or female at a given time while other plants may be functionally bisexual. Tong (1998) conducted a study discussing the hybridization of *C. tovarii* and the other Capsicum species. The results showed that genetically *C. tovarii* is more closely related to the *C. baccatum* complex than with the *C. pubescens* complex or the *C. annuum* complex.

4 CONCLUSIONS

Morphological data are traditionally utilized as a basis for taxonomic classification or grouping of *Capsicum* species or any other organism. The genus *Capsicum* has a very complex taxonomy, and its circumscription into one species or another can vary considerably based on the characteristics of the leaves, flowers, and fruits. These variations are often factors related to the geographic and weather conditions where the plants grow. However, these morphological bases are not enough, and therefore, molecular data are necessitated to establish its phylogenetic relationship. A study revealed rpl20 and rpl32, followed by rpl36, clpP, and accD as divergent genes from the chloroplast plastome. The gene sequence of rpl20, which is considered the most divergent gene, was used as a basis to establish a molecular-based phylogenetic tree. *C. galapagoense* and *C. tovarii* are more evolutionary diverged species whereas *C. lycianthoides* is at the primitive end of the tree.

The study generated a phylogenetic tree based on the rpl20 gene only to serve as baseline data for future studies on the Capsicum genus together along with studies utilizing morphological data. In order to obtain a comparative and thus a more accurate relationship of the Capsicum genus, it is recommended to create a phylogenetic tree based on other divergent gene sequence, especially, the *rpl3* gene as well as *rpl36, clpP*, and *accD* among others. In addition, more divergent Capsicum species such as *C. galapagoense* and *Capsicum tovarii* require more studies to be conducted on since further research work is limited to the endemicity of the species.

ACKNOWLEDGEMENT

The authors of the study would like to extend their gratitude to the Department of Biology, the Science and Mathematics Education Department of the University of San Carlos, and the Department of Science and Technology.

REFERENCE

Chiang T–Y and Schaal BA. (2000). Molecular Evolution

p-ISSN: 2599-4875 e-ISSN: 2599-4980

and Phylogeny of the atpB-rbcL Spacer of Chloroplast DNA in the True Mosses. Genome 43: 417-426.

Clement CR, Cristo-Araújo M, d'Eeckenbrugge GC, Pereira AA. (2002). Origin and domestication of Native Amazonian crops. Diversity 2: 72-106.

D'Agostino, Nunzio, Tamburino, R., Cantarella, C., De Carluccio, V., Sannino, L., Cozzolino, S., Cardi, T. Nunzia, S. (2018). The Complete Plastome Sequences of Eleven *Capsicum* Genotypes: Insights into DNA Variation and Molecular Evolution. Retrieved from https://www.mdpi.com/journal/genes. doi:10.3390/genes9100503.

Eshbaugh, W.H., Smith, P.G., Nickrent, D. L. (1983). Capsicum Tovarii (Solanaceae), A New Species Of Pepper From Peru. Brittonia. Vol. 35

FAOSTAT (2010). Agricultural Production Data. Available at [http://faostat.fao.org/]. Accessed May 14, 2012.

Jo, Y.D., Park, J., Kim, J., Song, W., Hur, C.-G., Lee, Y.-H., Kang, B.-C. (2011). Complete Sequencing and Comparative Analyses of the Pepper (*Capsicum annuum L.*) Plastome Revealed High Frequency of Tandem Repeats and Large Insertion/Deletions on Pepper Plastome. Plant Cell Rep. 30, 217–229.

Knapp, S. (2002). Floral diversity and evolution in the Solanaceae. In: Cronk QCB, Bateman RM and Hawkins JA (eds.) Developmental genetics and plant evolution. Taylor and Francis, London, p.267-297.

Kumar S , Stecher G, Tamura K. (2016). MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol Biol Evol . 33(7):1870–1874.

Miz RB, Mentz LA and Souza-Chies TT. (2008). Overview of the Phylogenetic Relationships of Some Southern Brazilian Species from Section Torva and Related Sections of "Spiny Solanum" (*Solanum subgenus Leptostemonum*, Solanaceae). Genetica 132: 143-158.

National Center for Biotechnology Information (NCBI). Accessed through https://www.ncbi.nlm.nih.gov/ Nei, M. and Kumar, S. (2000). Molecular Evolution and Phylogenetics. Oxford University Press, New York.

Passarin, D.M.M., Berger, I. J., Dressano, K., De Martin, V.F., Oliveira, G.C.X., Bock, R., and Carrer, H. (2008). Phylogenetic Relationships in Solanaceae and Related Species Based on cpDNA Sequence from Plastid trnEtrnT Region.

Reifschneider, F.J.B., Ribeiro, C.S.C. (2008). Cultivo. In: Pimentas *Capsicum* (Ribeiro CSC, Lopes CA, Carvalho SIC, Henz GM, et al., eds.). Embrapa Hortaliças, Brasilia, 11-14.

Saitou, N. and Nei, M. (1987). The Neighbor-joining Method: A New Method for Reconstructing Phylogenetic Trees. Molecular Biology and Evolution 4:406-425.

Tong, N. and Bosland, P. (1998). *Capsicum tovarii*, a New Member of the *Capsicum baccatum* complex. Department of Agronomy and Horticulture, New Mexico State University, Las Cruces, New Mexico, 88003, U.S.A.

Woodland, D.W. (1997). Contemporary plant systematics, 2nd ed., Andrews University Press, Barron Springs, Michigan. 619p.

p-ISSN: 2599-4875 e-ISSN: 2599-4980